Applicant : Jun Wu et al.
Serial No. : 10/802,479
Filed : March 16, 2004
Page : 14 of 20

Attorney's Docket No.: 16113-0615001 / GP-279-00-US

REMARKS

I.   Objection to the Specification

Paragraph 31 of the specification was objected to as containing content that was allegedly unclear. In particular, the Office Action objected to the content "…group remaining new character strings … into 7 sets of new characters" as being unclear, "because the context lacks description/definition of what the 7 sets really are and/or what the criteria/categories of the sets are used for grouping."

Paragraph 31 has been amended without adding new matter to more distinctly recite the subject matter that was already disclosed. As originally filed, Paragraph 31 described that the new words analyzer can analyze all character strings or those with less than or equal to a threshold number of characters, and provides an example where the new words analyzer can analyze those character strings having seven or fewer characters. The disclosure recites that the new words analyzer can group the character strings to be analyzed according to the number of Chinese characters that the character string contains. In other words, the character strings can be placed into groups of strings based on the number of characters in the character string. Therefore, when character strings containing seven or fewer characters are to be analyzed, there will be seven different sets or groups of character strings (i.e., contiguous characters). The seven groups result because a separate group is formed for the character strings that have seven contiguous characters, six contiguous characters, five contiguous characters, four contiguous characters, three contiguous characters, two contiguous characters, and one character, respectively. Paragraph 31 has been amended to clearly recite the aspect of the original disclosure.

No new matter has been added because the disclosure for the reasons set forth above already supported this amendment. Accordingly, withdrawal of the objection is requested.

II.   Claims 1, 3-5, 7-10, 12-14 and 16-17 are Allowable over the Art of Record

Applicant : Jun Wu et al.   Attorney's Docket No.: 16113-0615001 / GP-279-00-US
Serial No. : 10/802,479
Filed   : March 16, 2004
Page   : 15 of 20

Claims 1, 3-5, 7-10, 12-14 and 16-17 are pending in the present application, and claims 1, 3, 4, 7-10, 12, 13, 16, and 17 are amended. No new matter has been added. Support for the amendments can be found, for example, in paragraph 40 of the specification. Reconsideration of the pending claims is requested.

Claims 1-3, 5, 9-12 and 14 presently stand rejected under 35 U.S.C. § 103(a) as being rendered obvious by U.S. Publication No. 2007/0118346 to Badino ("Badino") in view of U. S. Patent No. 7,165,019 to Lee et al. ("Lee"). Independent claims 1, 9, and 10 have been amended to more distinctly claim the subject matter and distinguish over Badino, Lee, and the other art of record.

i.   Claim 1 is Allowable Over the Art of Record

Claim 1 has been amended to incorporate features not taught by the art of record, including features similar to those of canceled claim 6. The Office Action rejected claim 6 as being rendered obvious by Badino in view of Lee, and further in view of Nie et al. ("Unknown Word Detection and Segmentation of Chinese Using Statistical and Heuristic Knowledge", communications of COLIPS, vol. 5. NO 1&2, DEC 1995, page 47-57) ("Nie"). Particularly, Claim 1 has been amended to recite, in part:

> for each unknown character string,
>
> determining a corresponding first frequency of occurrence for the unknown character string and a corresponding **second frequency of occurrence for each of the Chinese characters in the unknown character string;**
>
> comparing the first frequency of occurrence to the second frequency of occurrence to determine an information gain value;
>
> comparing the information gain value to a threshold;
>
> **identifying the character string as a new valid word when the information gain is greater than the threshold...**

Applicant : Jun Wu et al.   Attorney's Docket No.: 16113-0615001 / GP-279-00-US
Serial No. : 10/802,479
Filed   : March 16, 2004
Page   : 16 of 20

The Office Action concedes that <u>Badino</u> in view of <u>Lee</u> does not expressly disclose determination of a valid new character string based on a predetermined threshold. Office Action at 7. Therefore, because claim 1 requires "comparing the information gain value to a threshold; identifying the character string as a new valid word when the information gain is greater than the threshold," claim 1 is allowable over <u>Badino</u> and <u>Lee</u>.

Amended claim 1 also clearly distinguishes over <u>Nie</u> at least because <u>Nie</u> does not disclose "comparing the information gain value to a threshold; identifying the character string as a new valid word when the information gain is greater than the threshold." At most, <u>Nie</u> discloses identification of potential new words based on a non-overlapping n-gram frequency of occurrence. <u>Nie</u> at 52. To determine a non-overlapping n-gram frequency of occurrence, <u>Nie</u> determines a raw frequency of occurrence for the n-gram (e.g., 2-gram). The raw frequency of occurrence is then reduced by the frequency of occurrence of the n-gram as a constituent part of a larger n-gram (e.g., 3-gram). The result is the non-overlapping frequency of occurrence. For example, if "add" is the n-gram being examined, the frequency of occurrence of "add" would be reduced in <u>Nie</u> by the frequency of occurrence of longer n-grams that include "add," e.g., "addition," "additive," etc.

Once the non-overlapping frequency of occurrence is determined, <u>Nie</u> eliminates a portion of the potential new words based on a fallout measure and a precision measure. <u>Nie</u> describes the fallout measure as the ratio of the number of eliminated words to the number of eliminated n-grams. <u>Nie</u> at 52-53. <u>Nie</u> describes the precision as the ratio of the number of real words found to the number of remaining n-grams. The remaining words are considered new words. <u>Nie</u> at 52-53. Thus, new words are identified by <u>Nie</u> without **any** consideration of the frequency of occurrence of each character that comprises the n-gram. Only after words are

added to the dictionary is there any consideration of the frequency of occurrence of each of the component characters that comprise the n-gram. Nie at 55.

In contrast to Nie, claim 1 compares "the first frequency of occurrence for the unknown character string" to the "corresponding **second frequency of occurrence for each of the Chinese characters in the unknown character string**" to determine an information gain. In turn, claim 1 identifies "the character string as a new valid word when the information gain is greater than the threshold." Because Nie does not consider the frequency of occurrence of each component character to identify new words, it cannot be cited as disclosing this feature of claim 1. Accordingly, withdrawal of the rejection of claim 1, and all claims depending directly or indirectly therefrom, is requested.

ii.   Claim 9 is Allowable Over the Art of Record

Claim 9 has been amended to include features not taught by the art of record, including features similar to those of canceled claim 6. The Office Action rejected claim 6 as being rendered obvious by Badino in view of Lee, and further in view of Nie. Particularly, Claim 9 has been amended to recite, in part:

for each unknown character string,

determining a corresponding first frequency of occurrence for the unknown character string and a corresponding **second frequency of occurrence for each of the Chinese characters in the unknown character string**;

comparing the first frequency of occurrence to the second frequency of occurrence to determine an information gain value;

comparing the information gain value to a threshold;

identifying the character string as a new valid word when the information gain is greater than the threshold...

Applicant : Jun Wu et al.   Attorney's Docket No.: 16113-0615001 / GP-279-00-US
Serial No. : 10/802,479
Filed   : March 16, 2004
Page   : 18 of 20

The Office Action concedes that <u>Badino</u> in view of <u>Lee</u> does not expressly disclose determination of a valid new character string based on a predetermined threshold. Office Action at 7. Therefore, because claim 9 requires "comparing the information gain value to a threshold; identifying the character string as a new valid word when the information gain is greater than the threshold," claim 9 is allowable over <u>Badino</u> in view of <u>Lee</u>.

Amended claim 9 also clearly distinguishes over <u>Nie</u> at least because <u>Nie</u> does not disclose "comparing the information gain value to a threshold; identifying the character string as a new valid word when the information gain is greater than the threshold," for substantially the same reasons as those discussed in reference to claim 1. Accordingly, withdrawal of the rejection of claim 9, and all claims depending directly or indirectly therefrom, is requested.

iii.   Claim 10 is Allowable Over the Art of Record

Claim 10 has been amended to include features not taught by the art of record, including features similar to those of canceled claim 6. The Office Action rejected claim 6 as being rendered obvious by <u>Badino</u> in view of <u>Lee</u>, and further in view of <u>Nie</u>. Particularly, Claim 10 has been amended to recite, in part:

> a new word analyzer configured to determine a corresponding first frequency of occurrence for the unknown character string and a corresponding second frequency of occurrence for each of the Chinese characters in the unknown character string, **compare the first frequency of occurrence to the second frequency of occurrence to determine if the character string is a new valid word based on a threshold,** and add the new valid word to the Chinese dictionary to create the updated Chinese dictionary

Applicant : Jun Wu et al.
Serial No. : 10/802,479
Filed : March 16, 2004
Page : 19 of 20

Attorney's Docket No.: 16113-0615001 / GP-279-00-US

The Office Action concedes that <u>Badino</u> in view of <u>Lee</u> does not expressly disclose determination of a valid new character string based on a predetermined threshold. Office Action at 7. Therefore, because claim 10 requires a new word analyzer that is configured to "determine a corresponding first frequency of occurrence for the unknown character string and a corresponding second frequency of occurrence for each of the Chinese characters in the unknown character string" and "compare the first frequency of occurrence to the second frequency of occurrence to determine if the character string **is a new valid word based on a threshold,**" claim 1 is allowable over <u>Badino</u> and <u>Lee</u>.

Amended claim 10 also clearly distinguishes over <u>Nie</u> at least because <u>Nie</u> does not disclose "compare the first frequency of occurrence to the second frequency of occurrence to determine if the character string is a new valid word based on a threshold," for substantially the same reasons as those discussed with reference to claim 1. Accordingly, withdrawal of the rejection of claim 10 and all claims depending directly or indirectly therefrom, is requested.

CONCLUSION

The allowability of all of the pending claims has been addressed. The absence of a reply to a specific rejection, issue, or comment does not signify agreement with or concession of that rejection, issue, or comment. In addition, because the arguments made above may not be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment or cancellation of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment or cancellation.

Please apply any charges or credits to deposit account 06-1050.

Applicant : Jun Wu et al.
Serial No. : 10/802,479
Filed : March 16, 2004
Page : 20 of 20

Attorney's Docket No.: 16113-0615001 / GP-279-00-US

Respectfully submitted,

Date: 7/1/08

Tracy M. Hitt
Reg. No. 62,371

1180 Peachtree Street, N.E., 21st Floor
Atlanta, GA 30309
Telephone: (404) 724-2775
Facsimile: (404) 892-5002

12060439.doc